RESEARCH PAPER

# Diabetes Mellitus Detection: A Comprehensive Comparative Analysis of traditional machine Learning, ensemble Approaches, and their Performance

**Poonam Sengar[1], Dr. Sandip Panchal[2]**

[1]Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India
[2]Vice-Principal, Dr. Subhash University, Junagadh, Gujarat

**ABSTRACT**
Diabetes Mellitus (DM) is a long-lasting metabolic disorder that disrupts the “body’s ability to maintain stable blood glucose levels” and has serious adverse effects on multiple organs. It leads to life-threatening complications like neuropathy, total kidney damage, and blindness. If DM can be detected early, patients will be able to avoid developing long-term complications. In the past few years, “Machine Learning (ML)” has become an increasingly effective method of clinical decision support by allowing the automated risk assessment, early disease detection, and faster diagnosis of patients’ health status. In this research, the main objective was to develop more precise methods of predicting diabetes by comparing traditional machine learning (ML) classifiers among Ensemble ML methods using the “PIMA Indian Diabetes Dataset”. We implemented a structured methodology performing pre-processing activities on the data. The performance of models has been assessed by means of “a variety of metrics. Experimental results demonstrate that Ensemble Models for predicting diabetes is much more successful “compared to Traditional Machine Learning Algorithms. Random Forest had the highest accuracy of 84%, XGBoost at 83%, and AdaBoost at 81%”. These results suggest that Ensemble Learning techniques provide more Accurate predictions.
**Keywords:** Accuracy, AdaBoost, Diabetes Detection, Kappa, Machine learning, and XGBoost.
**How to cite this article:** Sengar P, Panchal S. Diabetes Mellitus Detection: A Comprehensive Comparative Analysis of traditional machine Learning, ensemble Approaches, and their Performance. Int J Drug Deliv Technol. 2026;16(51s): 1629-1646. DOI: 10.25258/ijddt.16.51s.124

**Source of support:** Nil
**Conflict of interest:** None

## 1.INTRODUCTION

Diabetes mellitus is one of the fastest rising worldwide health issues of the 21st century. Diabetes mellitus is a metabolic condition in which the body develops resistance to the impact of insulin or does not produce enough of this hormone to digest glucose. In turn, there is a build-up of “sugar or glucose in the body, which can lead to major health concerns” [1-3]. “Diabetes affects persons of all ages, genders, and geographic areas, making it one of the most prominent global causes of mortality and morbidity. Type-1, Type-2, and GDM diabetes are the most prevalent [4]”. “Type-1 diabetes is often referred to as insulin-dependent DM. This is dubbed juvenile-onset DM, as it frequently occurs in infancy. Diabetes with type 1 is an autoimmune condition. This occurs impaired. This type of diabetes can be brought on by body factors. This may also develop due to difficulties with cells in the pancreas that make insulin. Non-insulin-dependent or adult-onset diabetes is another name for type 2 diabetes. Yet it has become frequent in children and teenagers in the last 20 years, especially since more young people are obese or overweight. When you have type-2 diabetes, your pancreas often releases some insulin”. However, either that is insufficient, or the body does not use it to its full potential. Type-2 DM is also comparatively milder than type-1 diabetes [5-7]. One of the most frequent problems seen in pregnant women is gestational diabetes mellitus (GDM). “Preterm labour, pre-eclampsia, nephropathy, delivery trauma, caesarean section, and later poor wound healing” are all possible outcomes of GDM. Classifications of Diabetes Mellitus are shown in Fig.1. GDM also increases the risk to the mother’s and the fetus’s health [8, 9].

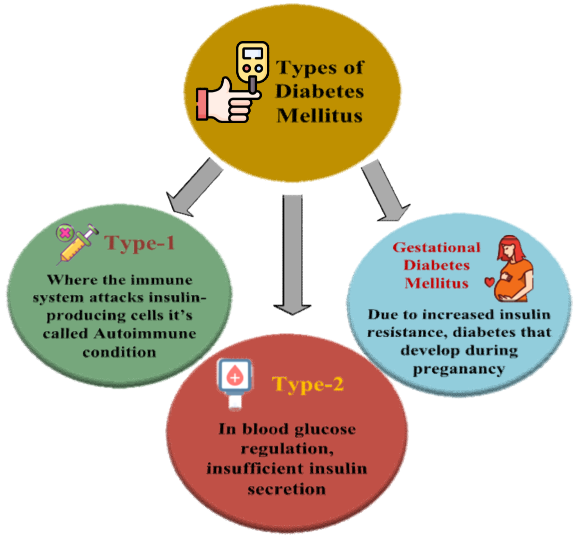


**Fig.1** Classification of Diabetes Mellitus

"Blood Sugar levels rise due to insufficient insulin production by the body", which can result in hyperglycemia and diabetes. Over time, this can seriously harm the human body and "lead to heart disease, stroke, kidney failure, blindness and amputation". Diabetes is referred to be the "second killer" among contemporary illnesses since its incidence is second only to that of cancer. Fig.2 represents the complications of diabetes mellitus. By 2045, 700 million persons worldwide will have diabetes, "according to the International Diabetes Federation". As a result, early disease detection allows patients who are at risk to initiate preventative measures to decrease disease development and improve quality of life [10, 11].

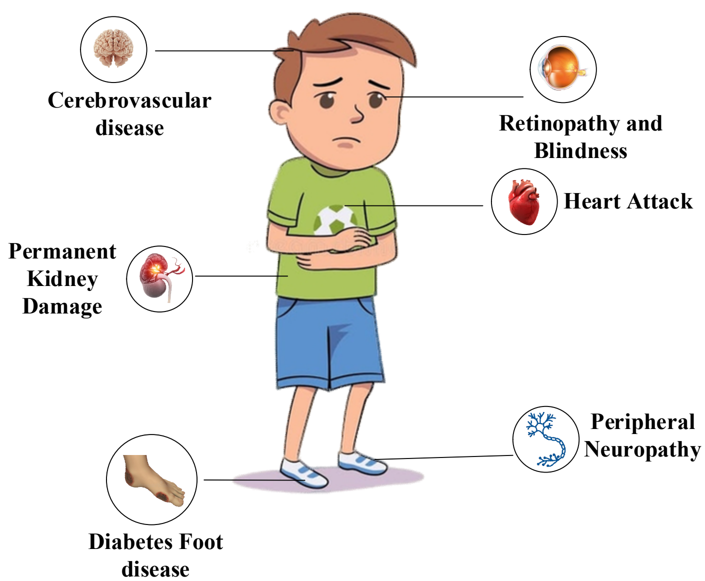


**Fig.2** Complications of Diabetes Mellitus

ML will help practitioners with future research and allow patients to easily confirm their health at an initial stage. "Early detection and diagnosis of type-2 diabetes" is now essential for medical science. Researchers have suggested and encouraged the "use of various ML techniques in the medical field as part of efforts to improve healthcare for diabetic patients. These techniques have a significant impact on type-2 diabetes early diagnosis, prediction, and prevention interventions. Various ML techniques for predicting type-2 diabetes include Decision Tree (DT), RF, Naïve Bayes (NB), Logistic Regression (LR), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM). It can be applied to problems involving both classification and regression". The categorization of people based on whether they have diabetes or not, since diabetes prognosis is a classification problem. Many "machine learning techniques" are useful for evaluating a classification problem [12, 13].

In the context, the Motivation and Contribution are as follows,

- Early detection is most in need of diabetes, and past research shows unreliable performance due to imperfect model comparison or limited pre-processing.
- A uniform process for pre-processing was established by using a pipeline for missing values and outliers, normalization, and feature extraction.
- Traditional and ensemble methods of machine learning will be evaluated under rigorous, controlled experimental conditions, giving identical comparisons of the models.
- The reliability of the models will be assessed using numerous assessment metrics such as "accuracy, precision, recall, f1-score, AUC, kappa, MAE, and RMSE".
- The results indicate that the most reliable models for the early prediction of Diabetes are the Ensemble model, with "RF and XGBoost" being the maximum performing models.

This paper is arranged as follows: Existing literature on "diabetes prediction using machine learning and ensemble learning" methods is reviewed in Section 2. Section 3 defines the proposed work. The performance comparison between the conventional ML and ensemble models is described in Section 4. Section 5 summarizes the key challenges of this study. And, finally, the conclusions of the work are presented in Section 6.

## 2. RELATED WORKS

This section discusses previous studies that have researched Diabetes Prediction through traditional ML and Ensemble learning approaches. Each of the studies that were reviewed has used "Logistic Regression, SVM, Decision Trees, KNN, Naïve Bayes classifier, RF, AdaBoost, or XGBoost". The majority of the studies reviewed showed that when comparing Ensemble to Single ML models, Ensemble Models generally outperform Single ML models; however, in most studies, consistent Pre-processing was not used, most studies do not compare multiple types of algorithms, and few studies used all performance metrics. The above illustrates that there is a need for a unified approach to compare and evaluate diabetes detection algorithms.

Authors in [14] propose "a novel ensemble learning-based framework" for the use of lifestyle indicators in the early prediction of Type-2 diabetes mellitus. Several ensemble learning strategies are used, including voting, bagging, and boosting. To enhance the dataset's quality assessment, exploratory data analysis is employed. Despite the system's alerts for potential diseased patients, the system is not supported to validate its ability to classify patients in the early stages. This work [15], they suggest "a machine learning model to help in diabetes prediction. LnR, LR, KNN, NB, (RF), SVM, and DT are some of the machine learning techniques included in the model". Although it remains less than the existing models. In this article [16], a very sensitive biosensor is used to detect diabetes using a "medical deep learning model". The invention of this extremely sensitive biosensor represents a major advancement in the detection of diabetes. It will help "enhance the accuracy and efficacy of diabetes diagnosis, allowing for better treatment of the illness". The effectiveness of detection relies on accuracy and speed limitations with the model, and as such, can impact the completeness of the diagnostic assessment.

This research [17] proposes effective frameworks for predicting diabetes mellitus based on limited data from women aged 21-81. The proposed frameworks include preliminary stages such as "data augmentation, attribute analysis, and missing data imputation". They employed Shapley Additive Explanation (SHAP) to identify key "features for fitting Extra Tree (ET), RF AdaBoost, and XGBoost models". Though Boosted tree models, like XGBoost and AdaBoost, are complex with respect to the SHAP method. Thus, boosted trees are likely to impede clinical decision-making in a transparent manner. In order to classify initial stage of diabetes patients, they provide "a fine-tuned RF algorithm with the best parameters (RFWBP) that is utilized in conjunction with the RF algorithm and feature engineering". They use the dataset to train both the suggested "RFWBP and traditional techniques, such as the AdaBoost algorithm, support vector machine, logistic regression, naïve bayes, multilayer perception, and a regular RF". However, it can produce inaccurate insulin and blood glucose data, which makes it more difficult to identify diabetes [18].

This article [19] proposes using "machine learning ensemble methods to predict diabetes. The ensemble includes KNN, Naïve Bayes, RF AdaBoost, and a newly developed Light Gradient Boosting Machine. The proposed ensembles inherit LightGBM's detection capacity to improve accuracy. Under fivefold cross-validation, the suggested ensemble models outperform other recent models. The KNN, AdaBoost, and LightGBM models reach 90.76% detection accuracy". However, Class imbalance issues were not fully addressed. In this article, [20] presents "an unsupervised cluster-based feature grouping" model for detecting initial stage of diabetes using an open-source dataset containing data from "520 diabetic cases". The highest Accuracy for the cluster-based and entire datasets is 99.57% and

99.03%, respectively. “Multi-layer perception (MLP), RF, and K-Nearest Neighbours (KNN) have the highest precision, recall, minimum mean square error (MSE), and F1-score (1.000). RF and support vector machine have 0.984, RF has 0.010, KNN has 0.067, and RF has 99.20% accuracy. Although the proposed work does not contain a user-friendly online application for real-time diabetes prognosis, it does not provide treatment recommendations.

In this research [21], they present “an ensemble univariate short-term predictive model of subcutaneous glucose concentration in T1D” with the goal of improving the error in the “hypoglycemic region”. As a result, the underlying basis functions are chosen to minimize the “percentage of erroneous predictions (EP) in the hypoglycemic region”, which is assessed using “continuous glucose error grid analysis (CG-EGA)”. The ensemble model’s basis models were chosen based on their superior performance in the “hypoglycemic range”. “The glucose predictive model’s performance is evaluated for each patient using three metrics: (i) pure error metrics (RMSE, MAPE, and MAE), (ii) Time Lag, which expresses the temporal delay between actual and predicted subcutaneous glucose concentrations, and (iii) CGEGA”, which evaluates the potential clinical impact of errors given the “glycaemic range”. The models will have less capacity for enhancing performance with respect to the various therapies that would be desired.

Authors in [22] propose a “robust framework for developing a diabetes prediction model to help with the clinical diagnosis of diabetes. Spearman correlation and polynomial regression are used in the framework for feature selection and missing value imputation”, correspondingly, from an approach that improves their respective performances. Additionally, our created “twice-growth deep neural network (2GDNN) model, the RF model, the support vector machine (SVM) model, and other supervised machine learning models are suggested for classification. Grid search and repeated stratified k-fold cross-validation are used to optimise the model's hyperparameters”. However, in this proposed work, both PIMA and LMCH datasets are limited in their use, as neither provides a comprehensive representation of every demographic, including children, non-pregnant women, and individuals of African descent.

In this research, [23] proposes a Hi-Le mode and a combination of highway and LeNet models for the purpose of early and precise diabetes identification. The “Hi-Le model outperforms its separate models in accuracy, F1-Score, precision, and recall, achieving 94% accuracy, 96% F1-Score, 94% precision, and 95% recall”. In order to identify and forecast diabetes in its early stages, they also suggested a “hybrid model called HiTCLe that combines Highway, Lenet, and a Temporal Convolutional Network (TCN)”. However, the synthetic sample produced by ProWSyn may not correctly characterize the population’s clinical forms, which could lead to overfitting with both Hi-Le and HiTCLe as examples of deep ensemble models. This article proposes [24] an “EL based on radial basis functions and Bayesian networks”. The suggested EL technique is “compared with the performance of five ML techniques: LR, DT classification, SVM, KNN, and RF”. The proposed model only compares conventional machine learning algorithms, while overlooking more advanced algorithms that could produce superior results.

This research [25] proposes a “diabetes prediction model based on a fused machine learning technique. The conceptual framework includes two types of models: Support Vector Machine (SVM) and Artificial Neural Network”. These models evaluate a “dataset to determine whether a diabetes diagnosis is positive or negative,” and this framework splits the data into “a 70:30 ratio between training and testing datasets”. Moreover, the overall negative diabetes cases are 87, but still the proposed model can predict 80 cases correctly, while it did not predict the remaining 7 cases. Authors in [26] propose a statistical-based prediction technique for data analysis. Initially, a dataset was created and stored using IPFS after data was gathered from IoT devices. The data will next be filtered and scaled using noise-invariant data expansion. Following the creation of an “adaptive RF technique to train data on the training dataset”, the suggested model was used to classify individuals with diabetes. The proposed model has limits because it has not been trained with comprehensive city-based datasets and lacks continuous monitoring and the capability of alerting hospitals and family practices.

This article [27] suggests a “machine learning-based method for predicting diabetes mellitus”. The proposed work compares and uses a number of “machine learning methods for diabetes prediction”, and ultimately identifies “the three ML classifiers with the highest accuracy: RF, GBM, and LGBM”. Two different kinds of datasets are used to determine the prediction accuracy. This proposed system is only evaluated against two data sets and is very reliant on data augmentation methods. This reliance means that the amount of variability represented in clinical practice may not be a true representation, thus hindering its desirability and ability to generalize to other settings. In this research, [28] uses a “hybrid deep learning model CNN-LSTM to classify diabetes using acceleration data from wearable sensors positioned on the hip, knees, and ankles”. “Two convolutional layers and two LSTM layers make up the suggested CNN-LSTM model”. CNN-LSTM can learn patterns for classification and extract significant features by merging the two models. “Accuracy, precision, recall, F1-score, AUC, and ROC are used to evaluate the performance of CNN-

LSTM models”. However, the proposed system does not include other types of signals that could improve its predictive capabilities and help detect diabetes-related gait abnormalities more accurately.

In this article [29], a “novel Ensemble Deep Learning (EDL) clinical decision support system for highly accurate diabetes prediction” is presented. The proposed EDL approach combines an ensemble learning-based stacking model with “Deep Learning (DL) architectures, such as Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNN)”. When the medical records are used in the evaluation of the proposed framework, it’s incomplete. This research [30] introduces a “dynamic k-NN model that modifies the ‘k’ value based on local data peculiarities, improving prediction accuracy. Two available datasets are used in this proposed approach, such as Breast Cancer Wisconsin (BCW) and PIMA Diabetes. They analyse our performance using some metrics, including precision, accuracy, recall, F1-score, and execution time”. Although the k-NN method produces precise predictions, it does not provide an explanation for a particular classification. This “lack of interpretability may hinder its adoption in clinical settings, where transparency is crucial”.

Authors in [31] propose an “e-diagnostic model for diabetes categorization using a machine learning algorithm, which may be implemented on the Internet of Medical Things (IoMT)”. This work evaluates PIDD and BRFSS to categorize diabetes. The proposed approach uses random oversampling to “balance classes, the interquartile range technique for outlier detection”, and the “Boruta algorithm to choose optimal features from datasets”. “The suggested approach uses some ML algorithms as RF, gradient boosting models, light gradient boosting classifiers, and decision trees,” for diabetes prediction. However, this proposed work evaluates only a single disease. In order to forecast diabetes in the “PIMA Indian Diabetes Dataset (PIDD)”, this work suggests “a lightweight artificial neural network (ANN) named Diabetic Lite (DBLite) and a novel attention-based loss function” [32]. Though it does not address privacy and interpretations of the models, as well as the scalability of the models in the real world.

This research [33] proposes “AHDHS Stacking, an ensemble learning framework for diabetes mellitus classification and diagnosis. It utilizes the harmony search (HS) algorithm and stacking, with two steps of feature selection and base-learner optimization. The model’s overall performance is improved by selecting features based on the average performance of all base learners”. An adaptive hyperparameter technique accelerates the iterative process. However, the proposed system used a dataset that contains only type 2 diabetes patients, which does not accurately reflect the circumstances of all diabetic patients. Table 1 provides a summary of the previous works.

**Table 1**

**Summary of Existing Works**

| Ref | Objective | Algorithms/Methods | Limitations |
|---|---|---|---|
| **[14]** | To predict early type-2 diabetes using an ensemble-based model. | Ensemble learning framework, bagging, boosting, voting | Small demographic groups are limited and clinical dataset validation is not considered. |
| **[15]** | For diabetes classification to compare the common ML models. | LR, SVM, KNN, DT, RF, NB | Accuracy is enough and incomplete pre-processing. |
| **[16]** | To detect diabetes using the deep learning with bio sensor. | Deep learning | Scalability is limited. |
| **[17]** | To enhance the diabetes prediction using feature engineering and ML. | RF, AdaBoost, XGBoost, Extra Trees | Still challenging to the model interpretability and reliability restricts for small datasets. |
| **[18]** | To predict diabetes system using optimized RF. | AdaBoost, RF, SVM, MLP, NB, LR | It can’t have perfect overfitting and robustness is limited. |
| **[19]** | To using modern ensemble techniques for enhanced diabetes detection. | RF, LightGBM, KNN, NB, AdaBoost | Insufficient Tuning process and incomplete class imbalance. |
| **[20]** | To build a cluster-based feature grouping model for diabetes prediction. | KNN, MLP, RF | Due to noisy data, clustering is not stable. |

| | | | |
|---|---|---|---|
| **[21]** | For hypoglycemic diabetic cases to design a short-term predictor. | Ensemble short-term predictive model | Hypoglycemic range is limited. |
| **[22]** | For correlation – based feature selection to create a robust ML framework. | SVM, 2GDNN, RF | Need for heavy pre-processing and specific datasets are restricted. |
| **[23]** | To improve diabetes identification introduces the hybrid deep learning. | CNN, Highway Networks, TCN | Due to deep architecture, overfitting is risk and misleading synthetic data. |
| **[24]** | For diabetes classification to develop the ensemble learning using RBF and Bayesian methods. | Bayesian Network, RBF Network | Highly non-linear data performance is poor. |
| **[25]** | To present a fused ML technique combines SVM and ANN. | SVM, ANN | Based on scaling features performance are varied and generalizability is limited. |
| **[26]** | To predict diabetes build an IoT-enabled pipeline. | Adaptive Random Forest. | Large-scale testing is limited. |
| **[27]** | For diabetes prediction to compare the GBM, LightGBM, and RF. | RF, LightGBM, GBM | Due to the unseen data performance are weak. |
| **[28]** | To use the Gait-based signals via DL for diabetes detection. | CNN-LSTM | It doesn't consider the clinical features. |
| **[29]** | For clinical decision support to build an ensemble deep learning. | CNN, ANN, LSTM | Low Interpretability. |
| **[30]** | For medical classification to suggest a dynamic KNN. | Dynamic KNN | Interpretability is low and sensitive performance based neighborhood choice. |
| **[31]** | To using feature ranking build an IoMT-based diagnostic model. | GBM, LightGBM, DT, RF | It only focuses single-disease so multiple dataset validation are lacked. |
| **[32]** | For diabetes detection to build a lightweight ANN model. | ANN | Scalability is limited and diverse dataset validation is lacked. |
| **[33]** | To create a stacking ensemble with two-stage feature selection. | Stacking Ensemble, Harmony Search | Due to diabetes types testing is limited. |

## 3. PROPOSED METHODOLOGY

This suggests work for "predicting diabetes utilizing ML and EL techniques" follows a clearly defined unified pipeline. The first step in this workflow is to obtain the PIMA diabetes dataset; the next step is to perform several pre-processing tasks that must be performed before utilizing this dataset in the model development phase. Pre-processing tasks include dealing with missing values, outlier removal via the IQR approach, normalizing the features in the dataset, and performing PCA on the dataset to extract the relevant features for the model development phase. After performing these pre-processing tasks, the PIMA diabetes dataset will be separated into two separate datasets so that an unbiased evaluation of the model can occur, Next, in controlled environment, we will implement several traditional models (LR, SVM, DT, Naïve Bayes, and KNN) as well as ensemble models(RF, XGBoost, and AdaBoost) for evaluation purposes. Hyperparameter tuning will then take place for each ensemble model for optimal performance during the model evaluation phase. Ultimately, all the models will be evaluated using multiple metrics to determine which technique is the best method for predicting diabetes at an initial stage. The overall proposed work architecture is shown in Fig.3.

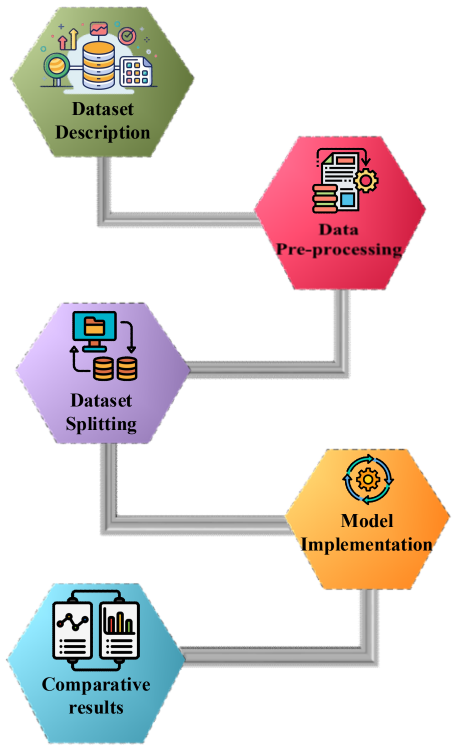


**Fig.3** Overall architecture of proposed work

### 3.1 Dataset description

In this study, the "PIMA Indian diabetes dataset (PID) was obtained from the Kaggle ML Repository. All of the patients in this study were adult women aged 21 or older [34]". The outcome was chosen as the aim. It contains information about 768 individuals who participated in the clinical study, as well as eight distinct parameters:

- Pregnancies
- Blood glucose
- Blood pressure
- Skin thickness
- Insulin
- Body mass index (BMI)
- Diabetic pedigree function (DPF)
- Age

### 3.2 Data pre-processing

"Data pre-processing" is a critical stage to make sure that predictive models are trained on reliable, meaningful inputs. For this study, we used the following pipeline on the dataset, which involves the following, multiple stages:

- ***Handling Missing Data:*** "Missing values are a prevalent problem in data analysis and machine learning", resulting from the absence of values for particular "variables or attributes during data collection or processing". "Missing values can lead to issues such as reduced sample size, information loss, and biased analysis results, potentially jeopardising the accuracy and reliability of data analyses and models [35] In order to minimize any potential data bias and guarantee that the data distribution is consistent with the normal distribution, we used a combined mean and median filling strategy to fill in the missing values".
- ***Outlier Detection*****:** To get reliable data analysis results, outliers-numbers in a dataset that differ significantly from the other data values and can negatively impact the distribution, linkages, and statistical analysis of the data-must be found and handled. For this feature to enhance the analysis quality, outlier handling is therefore necessary. Directly remove the outliers using the Interquartile Range method.
- ***Feature extraction*****:** The features with "numerical data (i.e., glucose levels, BMI) were normalized so that all attributes were on the same scale", improving the performance of the models for ML. Because of the increased normalization of features, the performance and accuracy of the model were improved. The most important components of the original dataset were extracted via "Principal Component Analysis (PCA) and used to reduce the dimensionality of the dataset" [36]. By extracting components with PCA, this procedure reduced the chance of multicollinearity among features while also improving the computation times.

- ***Dataset Imbalance*****:** Often, “hybrid sampling methods are used to address data imbalance issues seen in different classification techniques”. The common approach of doing this with using SMOTEENN combines SMOTE with ENN in order to increase the performance of Machine Learning models and rebalance the Dataset. The SMOTEENN method has two stages [35]. The first stage is performed by creating synthetic minority class samples through interpolation of “feature vectors between the minority class and its nearest neighbors”. In order to improve upon class distribution and create a more balanced dataset. Second, the merged dataset is subjected to the ENN method, which removes liar and ambiguous samples from both classes. By removing samples whose nearest neighbors misclassify them, ENN increases the Quality and Separability of the remaining dataset.

### 3.3 Dataset splitting

“PIMA Data set processing tasks to achieve clean & properly formatted input use by machine learning algorithm; it’s divided to the dataset into training and testing subsets”. “70% of the available data as a training set for ML models. The model was evaluated on unseen data using the Testing Set (30%)”. Splitting datasets correctly guarantees that the model performs well when applied to new, untested data. “Accuracy, precision, recall, F1-score, ROC-AUC, Kappa, MAE, and RMSE” were among the metrics used to compare the consequences of various “machine learning models and Ensemble models”.

### 3.4 Diabetes detection

“Traditional machine learning methods (LR, SVM, KNN, DT, and Naïve Bayes)” as well as ensemble methods (RF, AdaBoost, XGBoost) using a pre-processed data set. All models were trained under the same conditions to ensure that they could be compared fairly. The learning ability of the ensemble methods was improved through hyperparameter tuning. After completing the training process on the training data set, the models were estimated against a separate test data set using various metrics to assess each model’s capability to classify diabetic patients. Thus, all of the models were estimated expending similar metrics regarding “accuracy, precision, recall, F1-Score, Receiving Operating Characteristic (ROC) Area under the curve (AUC), Kappa coefficient, Mean Absolute Error (MAE), and Root-mean-square error (RMSE)” to ascertain which predictive modeling approach performed the best in identifying people with diabetes.

#### 3.4.1 Traditional machine learning models

Using “Machine Learning (ML)” is a data-driven approach that enables computers to learn from the history of the data used to make predictions, instead of having it programmed specifically. It detects diabetes; ML algorithms can use patients’ information (such as “glucose level, BMI, age, blood pressure, and insulin levels”) to determine if they are diabetic or not. In addition, ML provides faster and more consistent, and more automated methods of assessing an individual’s risk of disease.

***Support vector machine (SVM)****:* “One popular machine learning method” for classifying data points is “Support Vector Machine (SVM), which finds a hyperplane in an N-dimensional space”. Ruling the “hyperplane that maximizes the margin between various classes is the main goal of this algorithm”. The classifier’s robustness and generalization abilities are improved by this focus on maximizing the margin. The amount of “features determines the dimensionality, denoted by N. Comparing two features is very simple; however, handling several features for categorization might be more difficult. SVM improves prediction accuracy by optimizing the margin. SVM is an effective means of identifying Diabetes using “Supervised Classification” methods. By being able to represent the “Boundaries” between diabetic and non-diabetic patients using only the “Support Vectors”, SVM has excellent Predictive ability.

***Decision tree****:* “Decision Trees (DT) are predictive models that make use of a hierarchical tree structure [37]”. “In order to anticipate the goal value, which is found in the tree leaves, they arrange features, which are represented as branches”. “Classification trees are used when the target parameters have a limited range of values”. In these three models, the leaves represent class labels, while the branches show the feature specification that results in these class labels. It provides a prominent representation of how clinical professionals make diagnoses of patients at the time of Diagnosis and provides them with a Simple method for sorting Diabetic patients into various classes of diabetic patients.

***Naïve bayes****:* “Gaussian Naïve Bayes is a straightforward, highly functional, and computationally efficient classification algorithm [19]. It is a variation of the Naïve Bayes algorithm, which is applied when the features have continuous values and is based on the Bayes theorem with a Gaussian or normal distribution”. Naïve Bayes has demonstrated efficiency in many other medical studies, 0-2 samples. Naïve Bayes produces rapid results, which are extremely useful in the early diagnosis of diabetes.

***KNN****:* “KNN is a supervised learning”; a nonparametric classifier that uses K numbers of the closest “machine learning models” to approximate a function with discrete values [38]. In order to classify the data, it first creates a plane that has all of the training points on it, measures the distance between the query and the plane, and then generates a classification”. KNN compares patients to their nearest neighbors that present similar picture markers for the specified medical conditions in patients. If a cluster managed by Nearest Neighbor represents the structured cluster of relationships based on diabetes & Kidney failure or patients affected with both conditions, then this constitutes a strong Rationale for the identification or Diagnosis of Diabetes.

***Logistic regression**:* A statistical method for making predictions about binary classes is called logistic regression, or LR. It most closely resembles an "S-shaped function, which can be used to predict the outcome". LR offers a clear method for determining the "probability of being diagnosed as Diabetic based on each individual Clinical Feature" in order to assist health care professionals in identifying appropriate Risk factors. It also allows healthcare professionals access to information about individual Risk Factors to assist them in making diagnosis decisions.

### 3.4.2 Ensemble models

"Ensemble learning (EL)" utilizes advanced techniques of "combining the predictions of different models" to create an improved final prediction that contains greater accuracy than any individual model. "Ensemble methods such as RL, AdaBoost, and XGBoost" all improve upon the error rate and the degree to which they identify complex health patterns while simultaneously improving the stability of their predictions. Ensemble learning models, therefore, provide a more reliable result when it comes to making early diagnoses in diabetes detection.

***Random forest***: When it comes to ensemble techniques and hybrid models, RF is the maximum effective and versatile "supervised machine learning algorithm" for increasing performance and prediction accuracy. The RF technique uses bootstrapping to merge several decision trees into one ensemble model [19]. RF is extremely reliable for diabetes prediction because it decreases overfitting, captures non-linear clinical patterns, and offers high accuracy. Additionally, RF has the ability to identify key features (such as glucose and BMI) that have an impact on the clinical interpretation of diabetes.

***AdaBoost (Adaptive Boosting):*** AdaBoost is an ensemble method. This classifier fits multiple clones of the classifier to the same dataset after first working on the original dataset [39]. The AdaBoost algorithm primarily modifies the data distribution according to how well the training samples are classified. The lower classifier then gets the updated weights from the modified data once all of the training classifiers have been combined. AdaBoost struggles with noisy and outlier-containing datasets, even though it is less likely to overfit. In cases that are more difficult or misclassified patient cases, it improves diabetes classification by giving maximum weight. This adjustment results in a "decrease in error as well as an increase in the likelihood of detecting borderline diabetic cases correctly".

***XGBoost (Extreme Gradient Boosting)***: "XGBoost is an ensemble machine learning strategy that uses decision trees and gradient boosting" to increase performance by correcting faults in each algorithm's output. Compared to other algorithms, it is a more inventive and astute way. XGBoost performs well with both small and large datasets, although it struggles with datasets with a high number of categorical variables. XGBoost is one of the best-performing models for detecting diabetes because it can handle noise, missing values, or complex interactions between various clinical factors. The model's accuracy, combined with low levels of error, makes it an ideal choice for early-detection systems of diabetes.

### 3.4.3 Hyperparameter tuning

Machine learning needs to optimize model hyperparameters, which are not learnt from data. Optimal hyperparameter change can enhance model performance, generalization, and efficiency. To optimize the model parameters, so used the GridsearchCV approach. Ensure the model performs optimally in these parameters [40].

### 3.5 Confusion matrix

A confusion matrix [41], an analytics and machine learning efficiency evaluation tool, can be used to assess the "accuracy and effectiveness of a classification model". It carefully observes how the algorithm's predictions compare to the actual ground truth facts. It's also called a "square matrix", which shows the "total number of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) for each group or category in a classification problem".

- **True positives (TP):** A "positive class" was appropriately forecast by the process.
- **True negatives (TN):** A "negative class" was effectively forecast by the process.
- **False Positives (FP):** "A false positive occurs when an algorithm predicts a positive class while the actual class is negative".
- **False Negative (FN):** "A false negative occurs when an algorithm didn't predict the negative class while the actual class is positive".

"Classification algorithms can use the confusion matrix" to evaluate the "performance metrics such as accuracy, recall, and precision," and this provides useful insights into the system performance.

## 4. COMPARATIVE RESULTS

This results section is to evaluate the summary of the results obtained by comparing "machine learning and ensemble learning on PIMA diabetes data using metrics (precision, accuracy, recall, F1-score, ROC-AUC, Cohen's Kappa, MAE, and RMSE)" for analysis. This will give insight into how well the various models performed in detail in predicting which patients were diabetic and which were not, and demonstrate the added value provided by applying ensemble techniques.

***Accuracy:*** The "accuracy of a prediction model is defined as the number of correct predictions out of the total number of predictions" produced by the model. For "diabetes detection, the accuracy of a prediction model indicates how accurately the classifier can classify patients into the diabetic and non-diabetic categories. Fig.4 represents the accuracy comparison of learning models.

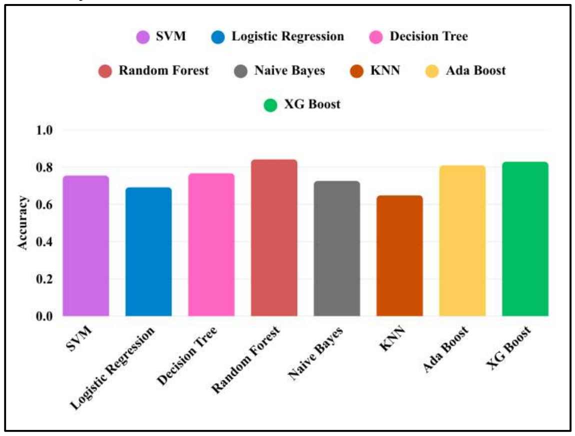


**Fig.4** Accuracy comparison of learning models

A higher accuracy will reduce the number of incorrect classifications by the model, ultimately reducing the number of errors within a clinical support system. In this graph, the "Random Forest [20] classifier" had the highest accuracy (84.18%), followed by the XGBoost (82.91%) and the AdaBoost classifiers (81.01%).

***Precision:*** The point to which a model accurately classifies those with diabetes from the entire dataset versus incorrectly classifying non-diabetes as diabetic is known as "precision". Therefore, the greater the indicated "precision", the fewer false positives will arise from the model's predictions. The proportion of correct positive predictions is shown in Fig.5.

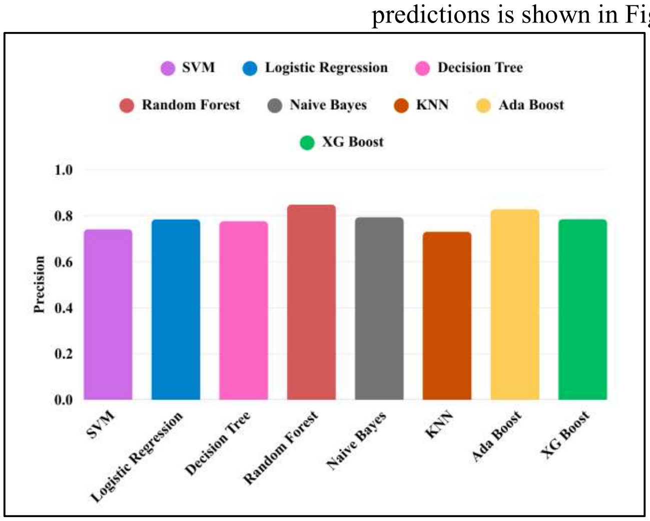


**Fig.5:** proportion of correct positive predictions

The graphs illustrate which models provide better reliability in identifying those diagnosed with diabetes. Typically, ensemble models sustain improved precision due to their boosted learning capability. The classifiers of RF (84.8%) and AdaBoost (83.7%) had the greatest precision values.

***Recall:*** To detect actual diabetic cases correctly, recall measures the model's ability. In the model, can missed few diabetes patients can represent the higher recall, which is critical in medical diagnosis. The higher recall aims to detect the most diabetic cases. The RF (83%) or XGBoost (87%) shows the superior recall since they are more accomplished at identifying the positive cases. Fig.6 represents the recall comparison across models.

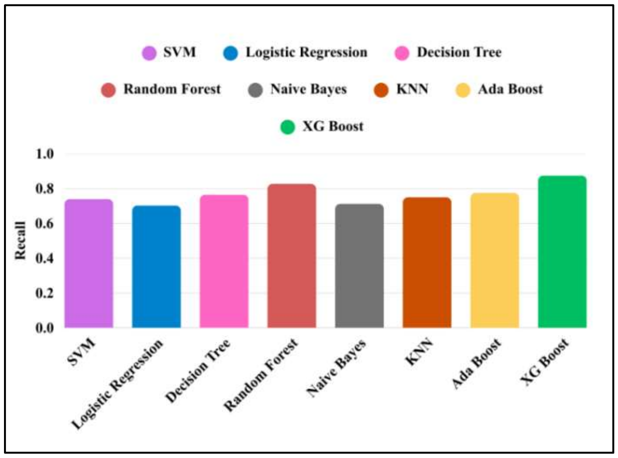


**Fig.6** Detection rate of actual diabetic cases

***F1-Score***: The "F1-score is a balance between both recall and precision"."F1-score is high", it demonstrates the performance of well-balanced prediction, and it establishes better stability in classifying diabetic and non-diabetic classes. F1-score comparison across models is shown in Fig.7.

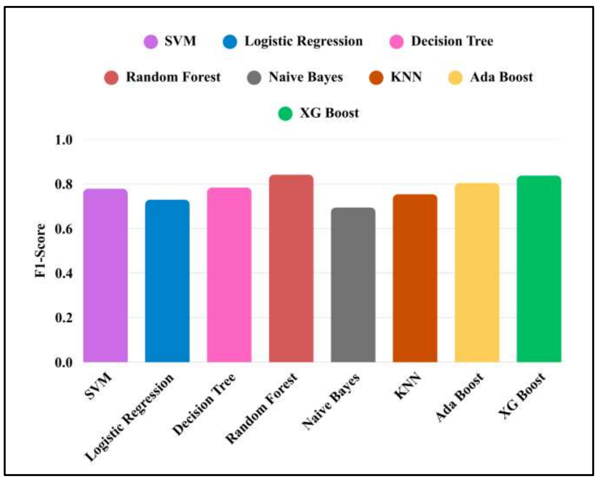


**Fig.7** F1-score comparison across models

In this graph, the RF (84.2%) and XGBoost (83.8%) had the highest value F1-score, so it does can create a better stability in classifying diabetes.

**Table 2**

**Comparison of Machine learning and Ensemble learning techniques using metrics Accuracy, precision, Recall, F1-score**

| Model | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|
| **SVM** | 0.78481 | 0.810811 | 0.75 | 0.779221 |
| **Logistic Regression** | 0.772152 | 0.814286 | 0.7125 | 0.76 |
| **Decision Tree** | 0.797468 | 0.815789 | 0.775 | 0.794872 |

| | | | | |
|---|---|---|---|---|
| **Random Forest** | 0.841772 | 0.848101 | 0.8375 | 0.842767 |
| **Naïve Bayes** | 0.765823 | 0.802817 | 0.7125 | 0.754967 |
| **KNN** | 0.778481 | 0.8 | 0.75 | 0.774194 |
| **AdaBoost** | 0.810127 | 0.837838 | 0.775 | 0.805195 |
| **XGBoost** | 0.829114 | 0.804598 | 0.875 | 0.838323 |

Performance of all Machine Learning and Ensemble methods compared across core metrics of classification, which consists of "accuracy, precision, recall, and F1-scores can be seen in Table 2. The consequences clearly show the highest accuracy when utilizing ensemble techniques over individual machine learning models. "RF provides the best overall accuracy (84.18%), and the F1-score is 84.27%. The highest recall was achieved by XGBoost. Logistic Regression, SVM, KNN, and Naïve Bayes are traditional Machine Learning models with moderate performance; however, none of these traditional models outperformed the ensemble models for predicting diabetes. Overall, this indicates that ensemble models produce the most accurate, sensitive, and balanced results for detecting diabetes.

***ROC-AUC:*** The model's ability to distinguish between "positive and negative classes is indicated by ROC-AUC". Strong discrimination ability is indicated by values nearer 1.0. According to the plot, ensemble models typically outperform conventional models in terms of AUC, representing their resilience in classification problems. The highest ROC-AUC is the best, so RF (92.8%) and XGBoost (91.3%) had the highest ROC-AUC values. Class separability performance of classifiers is shown in Fig.8.

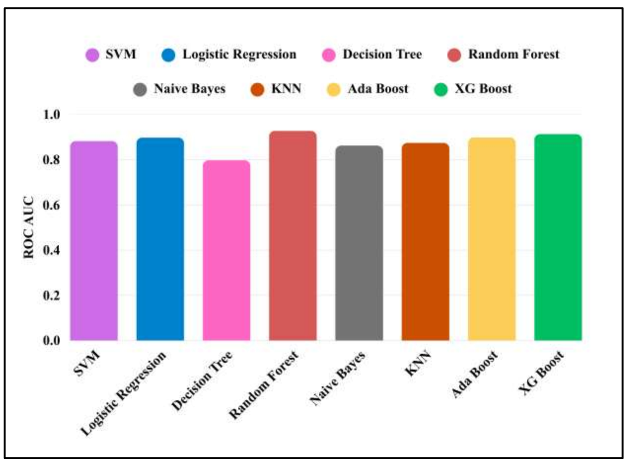


**Fig.8** Class separability performance of classifiers

***Kappa*:** This measure assesses the "degree of agreement between predictable and actual results while accounting for causal agreement". Reliability is better when kappa values are higher. Ensemble classifiers are typically RF (68.3%) and XGBoost (65.7%), exhibit higher Kappa scores, indicating reliable accuracy prediction. Fig.9 represents the agreement level between predictions and true labels.

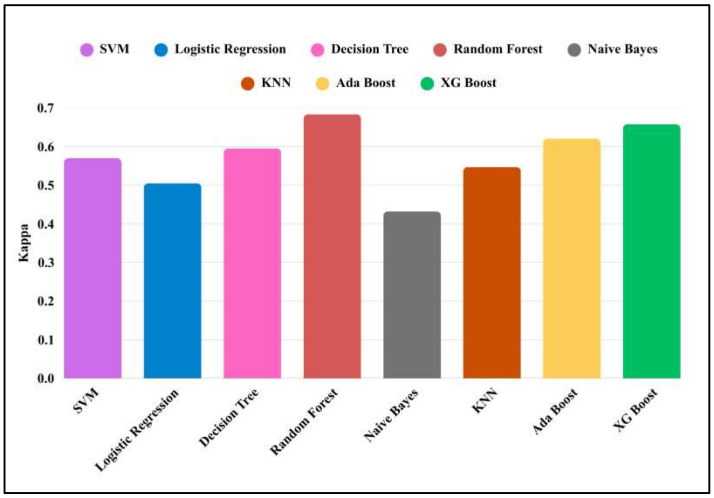


**Fig.9** Agreement level between predictions and true labels

***MAE:*** **"**Mean Absolute Error (MAE) estimates the average magnitude of prediction errors" without regard for direction. It is determined as the "average of the complete discrepancies between the actual and projected values". In this graph, the "values of Mean Absolute Error for ensemble learning and machine learning models" used for diabetes detection, RF (15.8%), and XGBoost (17%) can achieve the lower MAE. The Model comparison of MAE is shown in Fig.10.

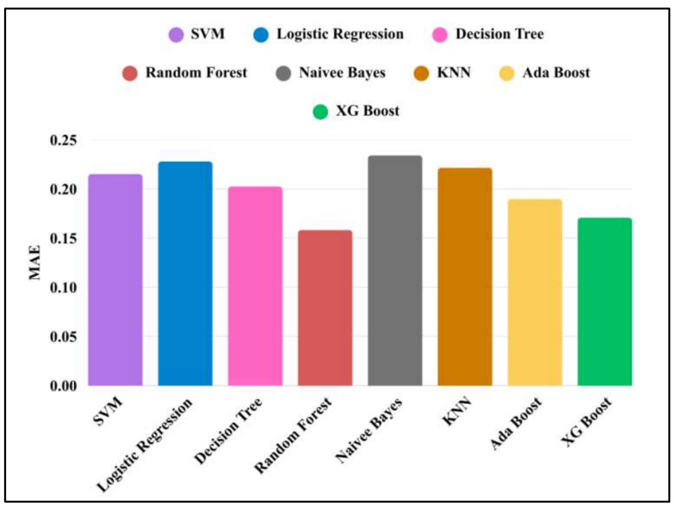


**Fig.10** Average absolute prediction error across models

***RMSE*****:** **"**Root Mean Squared Error (RMSE) metric" for assessing regression models. RMSE provides an estimate of the normal dissimilarity between forecast values and real values by capturing the "square root of the sum of the squared differences divided by the number of observations". Fig.11 explores the model comparison of RMSE.

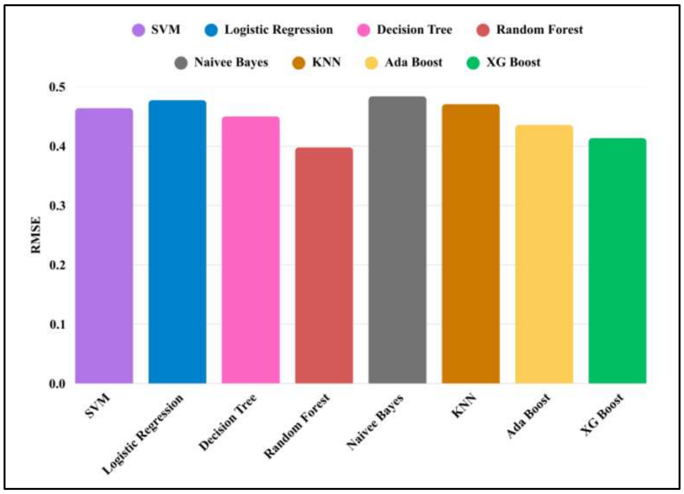


**Fig.11** Root-mean-square error comparison of models

For diabetes mellitus classifications, lower RMSE values indicate that the outputs of the model being evaluated are closer to the actual clinical values and therefore provide a more reliable way to "diagnose and early predict the risk of developing diabetes mellitus". So, this graph illustrates that the Ensemble learning techniques of RF (39.7%) and XGBoost (41.3%) have achieved the lower RMSE.

**Table 3**

**Comparison between the Machine learning and Ensemble learning techniques using metrics ROC AUC, Kappa, MAE, RMSE**

| Model | ROC AUC | Kappa | MAE | RMSE |
|---|---|---|---|---|
| **SVM** | 0.893029 | 0.569965 | 0.21519 | 0.463886 |
| **Logistic Regression** | 0.898397 | 0.54496 | 0.227848 | 0.477334 |
| **Decision Tree** | 0.797756 | 0.595131 | 0.202532 | 0.450035 |
| **Random Forest** | 0.928125 | 0.683544 | 0.158228 | 0.397779 |
| **Naïve Bayes** | 0.862821 | 0.532245 | .234177 | 0.483919 |
| **KNN** | 0.884936 | 0.557246 | 0.221519 | 0.470658 |
| **AdaBoost** | 0.898958 | 0.620557 | 0.189873 | 0.435745 |
| **XGBoost** | 0.913942 | 0.657789 | 0.170886 | 0.413384 |

Table 3 demonstrates how well the different learning models can predict outcomes using advanced performance measurements. Ensemble models such as RF(92%) and XGBoost (91%) as indicated by their high ROC-AUC scores which indicate very good class separation ability between Diabetic and Non-diabetic cases, their Kappa scores indicate very high agreement with actual labels of instances, Overall low MAE score can had the RF (15.8%) and XGBoost (17%), and lastly RF (39%) and XGBoost (41%) had the lowest RMSE scores, meaning they had less error in prediction and are therefore more stable. The traditional learning models had poor performance when compared to the ensemble models with lower AUC, kappa scores, and a higher degree of prediction error.

The overall "experimental results" indicate that the ensemble methods are superior to the traditional machine learning methods based on all of the evaluation criteria used in the study. The RF method was found to have the best accuracy rate (84.18%) and F1-score (84.27%). The XGBoost method produced the highest precision and recall (87.50 %). The results presented indicate that the "RF and XGBoost methods" are the most active way to diagnose diabetes and correctly identify those who have diabetes in less time than other methods available. Additionally, all classifiers classified as ensemble performed best out of all classifications with regard to "ROC-AUC, Kappa, MAE, and RMSE". Therefore, "compared to traditional methods, ensemble models" provided more accurate discrimination capabilities between "diabetes or non-diabetic patients".

Consistently, ensemble learning techniques outstripped "traditional machine learning models" in terms of minimizing errors and generalizing better to new data and stabilizing consistent classifications. It appears from our experimental results that Ensemble learning is a better approach than traditional ML techniques for diabetes prediction and provides the most accurate diagnostic assistance to allow for the diabetes detection for patient at initial stage.

## 5. DISCUSSION

Through the experimental data collection process and analysis, it was found that there was an important modification in how traditional ML algorithms performed in predicting diabetes as opposed to the ensemble learning algorithms. An examination of many different types of "metrics to evaluate each algorithm's performance (e.g., Accuracy, Precision, Recall, F1-Score, ROC-AUC, Kappa, Mean Absolute Error, and Mean Root Squared Error) revealed that all the ensemble learning models" produced consistently better and more robust results than traditional models. In particular, both RF and XGBoost exhibited excellent classification ability with greater than average accuracies and a more even distribution of sensitivity and specificity than that of Traditional ML models.

The traditional algorithms: SVM, KNN, LR, Naïve Bayes, and Decision Tree were able to provide a performance baseline that could be considered acceptable, but did not perform very well when it came to accurately modeling the complex interactions & non-linearities of the data. These traditional algorithms also tended to produce higher "mean absolute errors (MAE) and root mean square errors (RMSE)", which indicates that they performed worse at reducing misclassifications. The ensemble methods showed much lower error rates and the highest amount of agreement (kappa) among them, showing that these ensemble methods are much more reliable and appropriate for use in clinical prediction.

Ensemble learning has the capacity to lower the variability of classifications made by a classifier while increasing the capability of the classifier to more effectively differentiate between diabetic and non-diabetic patients through its ensemble mechanisms of bagging and boosting multiple weak learners together into an enhanced classifier. Furthermore, the recall and F1-Score metrics performed better with ensemble learning methods, as these metrics are extremely important in medical applications, where the failure to identify a diabetic patient can result in detrimental outcomes. Finally, Ensemble models provide a feasible option for diabetes early detection. By virtue of their ability to facilitate complex data management, reduce variance, and improve predictive reliability, they outperformed other methods in all the performance metrics examined. Consequently, integration of such models within the context of decision support systems used by health professionals will result in improved accuracy and efficiency of diagnosis.

## 6. CHALLENGES

Several challenges are encountered in diabetes detection, such as the following,

- The limited number of types from the "PIMA Diabetes dataset" makes it difficult to generalize; therefore, creating predictive models is restricted, and treatment decisions are not easy to interpret from clinical practice.
- Dimensionality Reduction of PCA provides a robust way of reducing dimensionality; however, new features would be challenging for clinical practice to understand.
- When using traditional ML models, the average performance was lower than ensemble-based models, indicating that traditional ML models have difficulty with recognizing and using complex patterns.
- Outlier detection for identifying and removing outliers from datasets needs caution, as excessive outlier removal can lead to a loss of valuable information.
- Low recall of the "Naïve Bayes classifier and Logistic regression" for diabetes indicates low capabilities of these classifiers to correctly identify true positive diabetes cases.

## 7. CONCLUSION

This section provides the concluding remarks of the work and suggests future enhancements to strengthen the "model's performance". In this conclusion, the research indicates that ensemble learning approaches tend to outperform conventional classifiers regarding their predictive ability for diabetes detection. Utilizing the "PIMA Indian Diabetes dataset for machine learning techniques" to "evaluate the performance metrics such as accuracy, precision, F1-score, recall, kappa, RMSE, and ROC-AUC". The findings reveal that "RF produced the greatest accuracy results at 84.18%, followed closely by

XGBoost 82.91% and AdaBoost 81.01%", thus providing further confirmation of the utility of ensemble learning methods to model complex healthcare data. In addition, it was notable that "XGBoost produced the highest recall score and RF produced the best F1-score", indicating that both models exhibited strong balanced performance. Based on these results, therefore, utilizing ensemble models could improve the reliability of a classifier's classification and reduce the probability of an incorrect diagnosis of diabetes, thus making them well-suited for use as the foundation for an intelligent system that will support the early diagnosis of diabetes. Future studies may consider using other advanced hybrid ensemble techniques and optimization strategies to "increase prediction accuracy of diabetes detection, utilizing hybrid deep learning architectures" and transformer-based models, in order to include a wider diversity of feature interactions from clinical datasets. Furthermore, it would be beneficial to expand the dataset by including actual clinical records and samples of patients treated across multiple sites to help "improve the model's generalization capabilities". Finally, the improvement of real-time predictive systems, such as mobile or IoT medical monitoring systems, can produce ongoing monitoring of diabetes risk and assist with prompt and accurate intervention.

**Author Contributions** P.S. conceptualized and designed the study and participated in data analysis and interpretation. Dr. S.P. assisted with data analysis and interpretation and provided critical feedback on the manuscript. All authors reviewed and approved the final version of the manuscript, and agreed to be responsible for all aspects of the work ensuring integrity and accuracy

**Funding**
This Research received no external funding.

**Data Availability**
The dataset used in this study is publicly available and can be accessed from the Kaggle repository.

## REFERENCES

1. Grundlingh, N., Zewotir, T. T., Roberts, D. J., & Manda, S. (2022). Assessment of prevalence and risk factors of diabetes and pre-diabetes in South Africa. Journal of Health, population and Nutrition, 41(1), 7.
2. Szkudelska, K., & Szkudelski, T. (2022). The anti-diabetic potential of betaine. Mechanisms of action in rodent models of type 2 diabetes. Biomedicine & Pharmacotherapy, 150, 112946.
3. Gollapalli, M., Alansari, A., Alkhorasani, H., Alsubaii, M., Sakloua, R., Alzahrani, R., ... & Albaker, W. (2022). A novel stacking ensemble for detecting three types of diabetes mellitus using a Saudi Arabian dataset: Pre-diabetes, T1DM, and T2DM. Computers in Biology and Medicine, 147, 105757.
4. Hossain, M. J., Al-Mamun, M., & Islam, M. R. (2024). Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused. Health Science Reports, 7(3), e2004.
5. Okyere, J., Ayebeng, C., & Dickson, K. S. (2024). Prevalence of diabetes and its associated factors in Cape Verde: an analysis of the 2020 WHO STEPS survey on non-communicable diseases risk factors. BMC Endocrine Disorders, 24(1), 264.
6. Chaki, J., Ganesh, S. T., Cidham, S. K., & Theertan, S. A. (2022). Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review. Journal of King Saud University-Computer and Information Sciences, 34(6), 3204-3225.
7. Ortiz-Martínez, M., González-González, M., Martagón, A. J., Hlavinka, V., Willson, R. C., & Rito-Palomares, M. (2022). Recent developments in biomarkers for diagnosis and screening of type 2 diabetes mellitus. Current diabetes reports, 22(3), 95-115.
8. Xiong, Y., Lin, L., Chen, Y., Salerno, S., Li, Y., Zeng, X., & Li, H. (2022). Prediction of gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning techniques. The journal of maternal-fetal & neonatal medicine, 35(13), 2457-2463.
9. Valadan, M., Bahramnezhad, Z., Golshahi, F., & Feizabad, E. (2022). The role of first-trimester HbA1c in the early detection of gestational diabetes. BMC pregnancy and childbirth, 22(1), 71.
10. Bhat, S. S., Selvam, V., Ansari, G. A., Ansari, M. D., & Rahman, M. H. (2022). Prevalence and early prediction of diabetes using machine learning in North Kashmir: a case study of district bandipora. Computational Intelligence and Neuroscience, 2022(1), 2789760.
11. Zhou, H., Xin, Y., & Li, S. (2023). A diabetes prediction model based on Boruta feature selection and ensemble learning. BMC bioinformatics, 24(1), 224.
12. Reza, M. S., Hafsha, U., Amin, R., Yasmin, R., & Ruhi, S. (2023). Improving SVM performance for type II diabetes prediction with an improved non-linear kernel: Insights from the PIMA dataset. Computer Methods and Programs in Biomedicine Update, 4, 100118.
13. NG, B. A. (2024). En-RfRsK: An ensemble machine learning technique for prognostication of diabetes mellitus. Egyptian Informatics Journal, 25, 100441.
14. Ganie, S. M., & Malik, M. B. (2022). An ensemble machine learning approach for predicting type-II diabetes mellitus based on lifestyle indicators. Healthcare Analytics, 2, 100092.

15. Uddin, M. A., Islam, M. M., Talukder, M. A., Hossain, M. A. A., Akhter, A., Aryal, S., & Muntaha, M. (2024). Machine learning based diabetes detection model for false negative reduction. Biomedical Materials & Devices, 2(1), 427-443.
16. Armghan, A., Logeshwaran, J., Sutharshan, S. M., Aliqab, K., Alsharari, M., & Patel, S. K. (2023). Design of biosensor for synchronized identification of diabetes using deep learning. Results in Engineering, 20, 101382.
17. Ejiyi, C. J., Qin, Z., Amos, J., Ejiyi, M. B., Nnani, A., Ejiyi, T. U., ... & Okpara, C. (2023). A robust predictive diagnosis model for diabetes mellitus using Shapley-incorporated machine learning algorithms. Healthcare Analytics, 3, 100166.
18. Ali, M. S., Islam, M. K., Das, A. A., Duranta, D. U. S., Haque, M. F., & Rahman, M. H. (2023). A novel approach for best parameters selection and feature engineering to analyze and detect diabetes: Machine learning insights. BioMed Research International, 2023(1), 8583210.
19. Sai, M. J., Chettri, P., Panigrahi, R., Garg, A., Bhoi, A. K., & Barsocchi, P. (2023). An ensemble of light gradient boosting machine and adaptive boosting for prediction of type-2 diabetes. International Journal of Computational Intelligence Systems, 16(1), 14.
20. Hassan, M. M., Mollick, S., & Yasmin, F. (2022). An unsupervised cluster-based feature grouping model for early diabetes detection. Healthcare Analytics, 2, 100112.
21. Katsarou, D. N., Georga, E. I., Christou, M. A., Christou, P. A., Tigas, S., Papaloukas, C., & Fotiadis, D. I. (2025). Optimizing hypoglycaemia prediction in type 1 diabetes with Ensemble Machine Learning modeling. BMC Medical Informatics and Decision Making, 25(1), 46.
22. Olisah, C. C., Smith, L., & Smith, M. (2022). Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective. Computer Methods and Programs in Biomedicine, 220, 106773.
23. Shaheen, I., Javaid, N., Alrajeh, N., Asim, Y., & Aslam, S. (2024). Hi-le and HiTCLe: Ensemble learning approaches for early diabetes detection using deep learning and explainable artificial intelligence. IEEE Access, 12, 66516-66538.
24. Mahesh, T. R., Kumar, D., Vinoth Kumar, V., Asghar, J., Mekcha Bazezew, B., Natarajan, R., & Vivek, V. (2022). Blended ensemble learning prediction model for strengthening diagnosis and treatment of chronic diabetes disease. Computational Intelligence and Neuroscience, 2022(1), 4451792.
25. Ahmed, U., Issa, G. F., Khan, M. A., Aftab, S., Khan, M. F., Said, R. A., ... & Ahmad, M. (2022). Prediction of diabetes empowered with fused machine learning. Ieee Access, 10, 8529-8538.
26. Azbeg, K., Boudhane, M., Ouchetto, O., & Jai Andaloussi, S. (2022). Diabetes emergency cases identification based on a statistical predictive model. Journal of Big Data, 9(1), 31.
27. Ahamed, B. S., Arya, M. S., & Nancy, A. O. V. (2022). Diabetes mellitus disease prediction using machine learning classifiers with oversampling and feature augmentation. Advances in Human-Computer Interaction, 2022(1), 9220560.
28. Chee, L. Z., Sivakumar, S., Lim, K. H., & Gopalai, A. A. (2024). Gait acceleration-based diabetes detection using hybrid deep learning. Biomedical Signal Processing and Control, 92, 105998.
29. Al Reshan, M. S., Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., Shaikh, A., & Elmagzoub, M. A. (2024). An innovative ensemble deep learning clinical decision support system for diabetes prediction. IEEE Access.
30. Al-Khamees, H. A., Sani, N. S., Gifal, A. S., Liu, L. X. W., & Esa, M. I. (2025). A dynamic model using k-NN algorithm for predicting diabetes and breast cancer. Computers in Biology and Medicine, 192, 110276.
31. Ashisha, G. R., Mary, X. A., Kanaga, E. G. M., Andrew, J., & Eunice, R. J. (2024). Random oversampling-based diabetes classification via machine learning algorithms. International Journal of Computational Intelligence Systems, 17(1), 270.
32. Roy, S., Cherish, R. R., & Roy, G. (2025). An attention-based loss function and synthetic minority oversampling technique for alleviating class imbalance in predicting diabetes. Healthcare Analytics, 100399.
33. Zhang, Z., Lu, Y., Ye, M., Huang, W., Jin, L., Zhang, G., ... & Zhu, W. (2024). A novel evolutionary ensemble prediction model using harmony search and stacking for diabetes diagnosis. Journal of King Saud University-Computer and Information Sciences, 36(1), 101873.
34. Patro, K. K., Allam, J. P., Sanapala, U., Marpu, C. K., Samee, N. A., Alabdulhafith, M., & Plawiak, P. (2023). An effective correlation-based data modeling framework for automatic diabetes prediction using machine and deep learning techniques. BMC bioinformatics, 24(1), 372.
35. Feng, X., Cai, Y., & Xin, R. (2023). Optimizing diabetes classification with a machine learning-based framework. BMC bioinformatics, 24(1), 428.
36. Ahmed, A., Khan, J., Arsalan, M., Ahmed, K., Shahat, A. A., Alhalmi, A., & Naaz, S. (2024, December). Machine learning algorithm-based prediction of diabetes among female population using pima dataset. In Healthcare (Vol. 13, No. 1, p. 37). MDPI.

37. Alnowaiser, K. (2024). Improving healthcare prediction of diabetic patients using KNN imputed features and tri-ensemble model. IEEE Access, 12, 16783-16793.
38. El-Sofany, H., El-Seoud, S. A., Karam, O. H., Abd El-Latif, Y. M., & Taj-Eddin, I. A. (2024). A proposed technique using machine learning for the prediction of diabetes disease through a mobile app. International Journal of Intelligent Systems, 2024(1), 6688934.
39. Tasin, I., Nabil, T. U., Islam, S., & Khan, R. (2023). Diabetes prediction using machine learning and explainable AI techniques. Healthcare technology letters, 10(1-2), 1-10.
40. Islam, M. M., Rifat, H. R., Shahid, M. S. B., Akhter, A., Uddin, M. A., & Uddin, K. M. M. (2025). Explainable Machine Learning for Efficient Diabetes Prediction Using Hyperparameter Tuning, SHAP Analysis, Partial Dependency, and LIME. Engineering Reports, 7(1), e13080.
41. Althobaiti, T., Althobaiti, S., & Selim, M. M. (2024). An optimized diabetes mellitus detection model for improved prediction of accuracy and clinical decision-making. Alexandria Engineering Journal, 94, 311-324.